

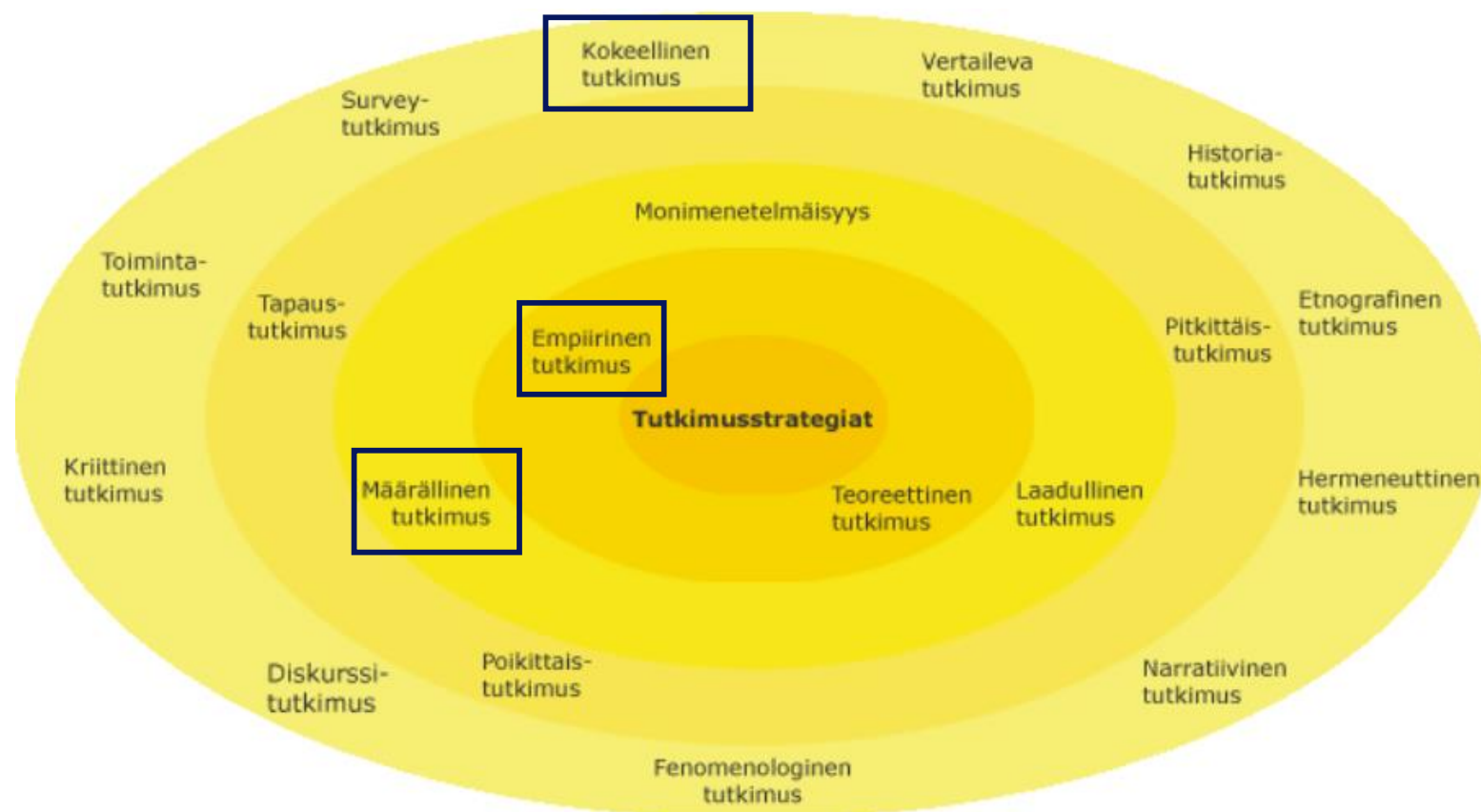
Määrällisen tutkimuksen ensiaskeleet – Otoskoko ja voima

Biostatistiikan neuvontapalvelut, Tutkimuspalveluyksikkö, Pohde

Biostatistikko Marko Merikukka (+ neuvontapalveluiden muut biostatistikot)

3.5.2023

Viitekehys



Otos vs näyte

Perusjoukko

Otos

- Perusjoukko (*population*) on kaikkien havaintoyksiköiden muodostama joukko
 - Esim. Kaikki Suomessa asuvat henkilöt
- **Otos** on satunnaisesti valittu havaintoyksikköjoukko
 - Esim. Puoluekannatusmittaus suomalaisilta
- Jos satunnainen havaintoyksikköjen valinta ei ole mahdollista, puhutaan **näytteestä**
 - Esim. huumeita käyttävien haastattelut
 - Ei tiedetä huumeiden käyttäjien määrää

Otanta – tapa poimia otos 1/2

- ~~Mukavuusotos~~

Otoskoko laskettava

- Yksinkertainen satunnaisotanta (*simple random sampling*)
- Systemaattinen satunnaisotanta (*systematic sampling*)
- Ositettu otanta (*stratified sampling*)
 - Edustuksellisuus merkittävien ryhmien osalta
- Ryväotanta (*cluster sampling*)
 - Esimerkiksi otos sairaaloista ja sen jälkeen potilaista

Otanta – tapa poimia otos 2/2

Otoskokoa ei tarvitse laskea

- Kokonaisotanta
 - Esim. rekisteritutkimus, jossa kaikki tutkittavat ovat mukana
- Kartoittava kysely
 - Esim. sähköinen kysely Pohteen työntekijöille
 - vastausprosentti

Otoskoko 1/3

= Havaintoyksikköjen määrä

- Odotettu otoskoko
- Toteutunut otoskoko

- Otoskoon tulisi aina olla riittävän suuri, jotta tutkimuksella pystyttäisiin vastaamaan tutkimussuunnitelmassa esitettyihin tutkimuskysymyksiin

Otoskoko 2/3

- Otoskoon tulisi pohjautua puhtaasti tieteellisille vaatimuksille ja havaintoyksikön kannalta merkitykselliseen tulokseen
 - Esim. Ei pelkkä laboratorioarvojen lasku vaan minkä suuruisella laboratorioarvojen laskulla voidaan ehkäistä esim. sydäninfarktia

Otoskoko 3/3

- P -arvot ja luottamusvälit riippuvat otoskoosta
 - Suurempi otoskoko – kapeammat luottamusvälit ja pienempi p -arvo
 - Pienempi otoskoko – leveämmät luottamusvälit ja suurempi p -arvo
- Mikä on riittävä otoskoko, jotta kliinistesti merkitsevä ero voidaan todentaa aineistoilla
 - tuhlaamatta henkilöstö- ja talousresursseja
 - aiheuttamatta havaintoyksiköille (esim. potilaille) turhaa vaivaa?

Liian pieni tai liian suuri otoskoko ei ole järkevää

→ Tehtävä voimalaskelmat (*power calculations*) (myös otoskokolaskelmat (*sample size calculation*) tai voima-analyysi (*power analysis*))

Otoskoko ja voimalaskelmat 1/2

Riittävän otoskoon laskemiseksi voimalaskelmiin tarvitaan

1. *Kliinisesti merkitsevä ero* ryhmien välillä!

- *tilastollisesti merkitsevä ero* ei aina ole kliinisesti merkitsevä
- $p_1 - p_2$ (osuuksien erotus, *kvalitatiiviset voimalaskelmat*)
- $x_1 - x_2$ (keskiarvojen erotus, *kvantitatiiviset voimalaskelmat*)
- tehtävä päätös substanssiosaamisen perusteella
- muutoin voimalaskelmat ovat kaavoihin sijoittamista

2. Päävasteen esiintyvyys tutkimusjoukossa

- päävasteen hajonta estimoidaan aikaisempien tutkimusten perusteella
- tarvittaessa tehtävä pilottitutkimus (ei tarvitse laskea otoskokoa)

Otoskoko ja voimalaskelmat 2/2

3. Merkitsevyystaso

- tyypillisesti $\alpha = 0,05$ (α = tyypin I virhe))

4. Tilastollinen voima (*statistical power*)

- tyypillisesti 0,8 tai 0,9
 - todennäköisyys havaita tilastollisesti merkitsevä ero, jos tietyn suuruinen ero ryhmien välillä on todella olemassa
 - jos voima on 80 %: samaa koetta lukuisia kertoja peräkkäin tehtäessä, neljässä kokeessa viidestä havaitaan tilastollisesti merkitsevä ero, jos se on olemassa
- = $(1 - \beta) \times 100$, missä β = tyypin II virhe ($0,1 \leq \beta \leq 0,2$)

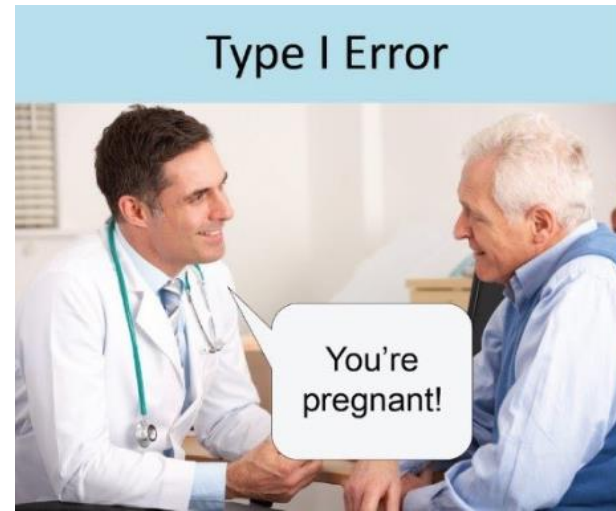
Tyypin I ja II virheet

α – tyypin I virhe (*false positive*)

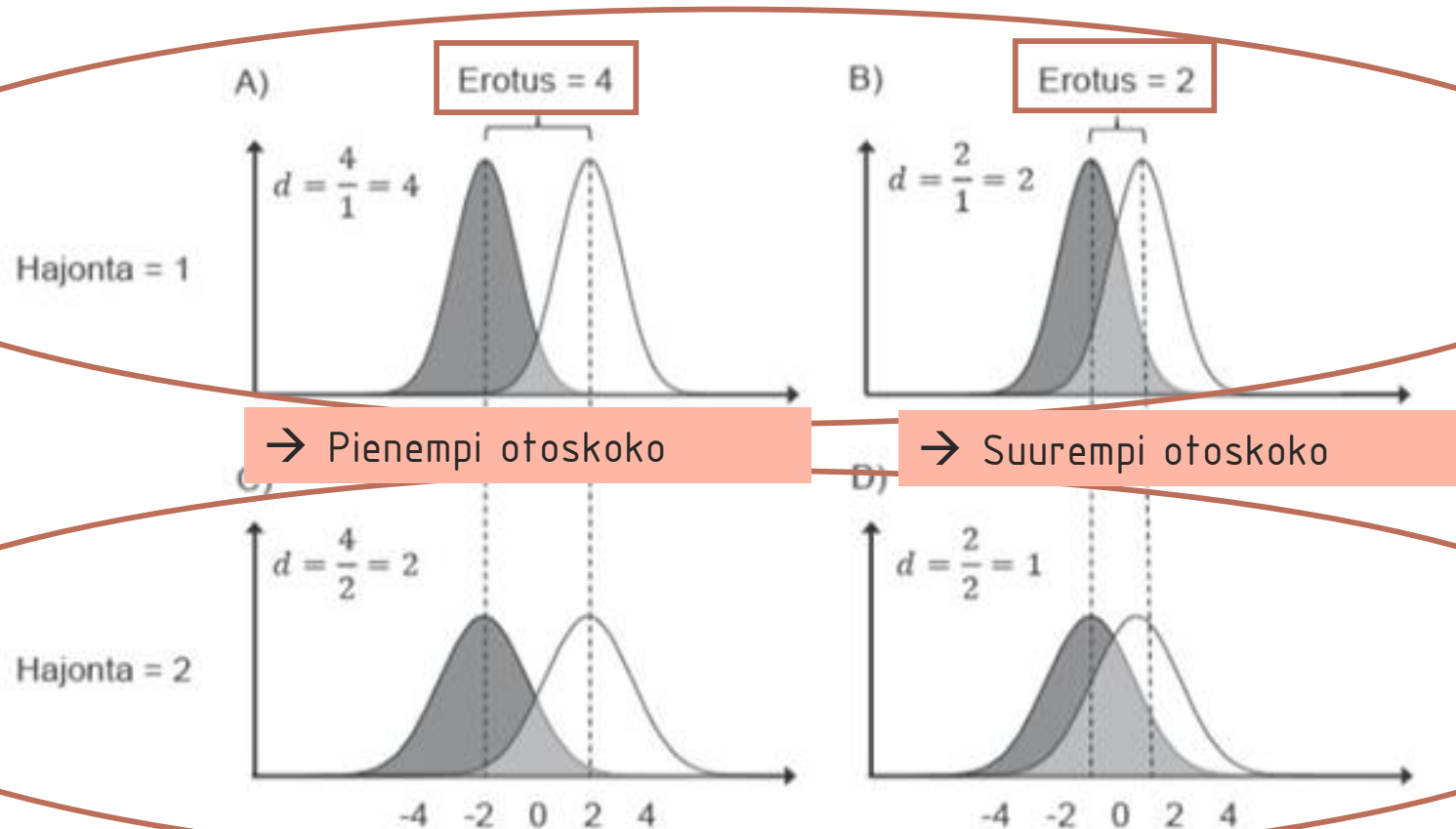
- Todennäköisyys havaita tilastollisesti merkitsevä tulos, vaikka nollahypoteesi H_0 pitäisi paikkansa (ero johtuu sattumasta 5 % / joka 20. kerta)
- H_0 : ryhmien välillä ei ole eroa vasteen arvoissa

β – tyypin II virhe (*false negative*)

- Todennäköisyys havaita tilastollisesti ei-merkitsevä tulos, vaikka vaihtoehtoinen hypoteesi H_1 pitäisi paikkansa (10–20 %)
- H_1 : ryhmien välillä on eroa vasteen arvoissa
- Riittämättömän otoskoon tyypillinen seuraus



Mitä pienempää kliinisesti merkitsevää eroa haetaan, sitä suurempi otoskoko tarvitaan



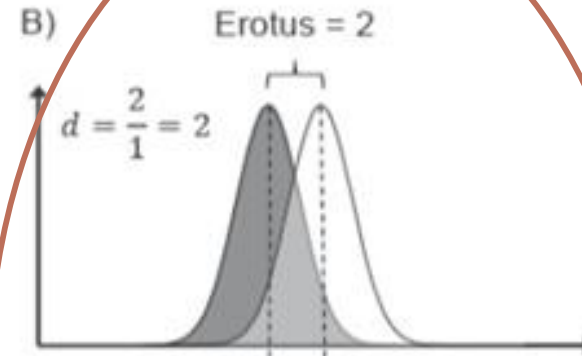
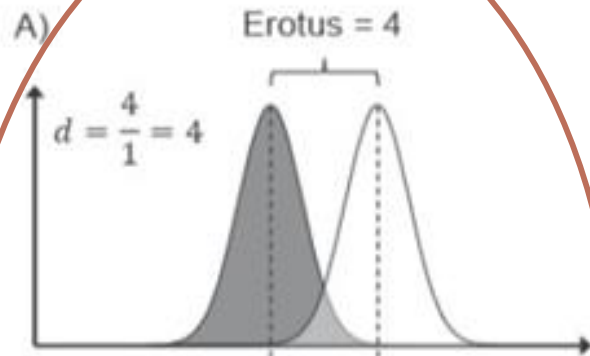
Jakaumien päällekkäin menevä massa on suurempi, mitä pienempi ero

Mitä pienempi vasteen hajonta, sitä pienempi otoskoko tarvitaan

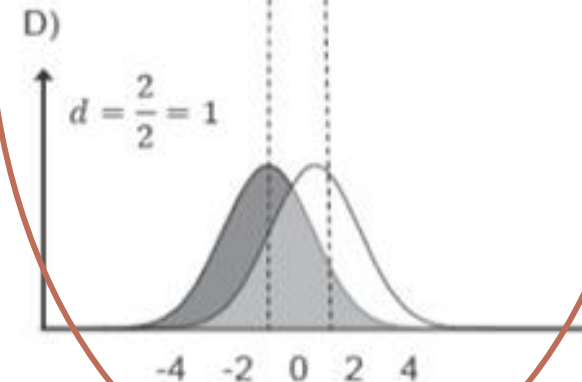
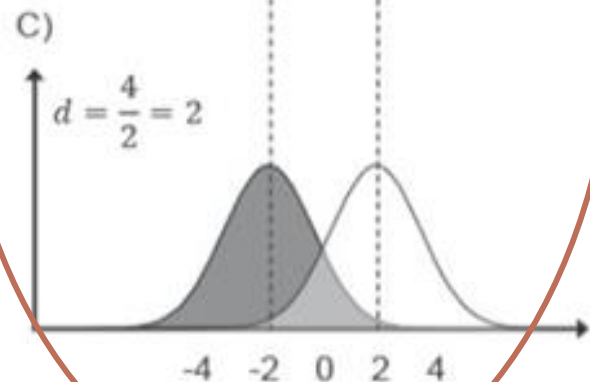
→ Pienempi otoskoko

Jakaumien päällekkäin menevä massa on pienempi, mitä pienempi hajonta

Hajonta = 1



Hajonta = 2



→ Suurempi otoskoko

Voimalaskelmat 1/3

1. Mikä on tutkimuksen päätarkoitus?
 - Tutkimuskysymys ja -hypoteesit
2. Mikä on tutkimuksen ensisijainen vaste?
3. Miten aineisto analysoidaan eron havaitsemiseksi?
 - Valittu tilastomenetelmä vaikuttaa voimalaskelmien kaavaan
4. Millaisia tuloksia on odotettavissa vertailuryhmässä?
 - Jotta voidaan laskea ero ryhmien välillä
5. Mikä on pienin kliinisesti merkitsevä ero, joka pitää havaita?

Voimalaskelmat 2/3

1. Tutkitaan onko uudelleensynnyttäjien lapset painavampia kuin ensisynnyttäjien lapset?
2. Ensisijainen vaste on lapsen syntymäpaino
3. Kahden riippumattoman ryhmän vertailu: riippumattomien ryhmien t -testi
4. Ensisynnyttäjien lapset painavat keskimäärin 3 400 grammaa. Hajonta oletetaan molemmissa ryhmissä olevan 500 grammaa aikaisemman kirjallisuuden perusteella.
5. Kliininen ero 200 grammaa

Lisäksi merkitsevyystaso on 0,05 ja tilastollinen voima 0,80

Biostatistikko tekee voimalaskelmat tilasto-ohjelmalla (esim. StatsDirect)



→ Otoskoko ryhmittäin 99 lasta

Voimalaskelmat 3/3

- Kato huomioitava voimalaskelmia tehdessä, jotta ryhmittäiset otoskoot ovat ylipäänsä mahdollisia
- Odotettu otoskoko ei riitä, otoskoon täytyy myös toteutua, jotta voimaa on riittävästi eron havaitsemiseksi
- Jos kadon oletetaan olevan 20 prosenttia, niin tarvittava otoskoko on 1,25-kertainen ($= N / (1 - 0,2)$)
- Oletetaan äskeisessä esimerkissä, että 10 prosentilla vastasyntyneistä ei saada mitattua painoa.
 - Korjattu otoskoko per ryhmä $= 99 / (1 - 0,1) = 110$

Voimalaskelmat tutkimuksen jälkeen?

- Otoksoon riittävyden arvioinnissa ei kannata turvautua voimalaskelmiin (vaikka vertaisarvioitsijat tätä usein pyytääkin)
- Otoksoon riittävyden tulisi perustua saatujen tulosten luottamusvälitarkasteluun
 - Luottamusvälin ala- ja ylärajan tarkastelu
- Luottamusväli ilmoittaa haarukan, jonka sisällä todellinen arvo tietyllä (yleensä 95 %) todennäköisyydellä sijaitsee

Tulosten vahvuuden karkea jaottelu

	Tulokset tilastollisesti merkitseviä	
Voimalaskelmat	Ei	Kyllä
Ei tehty	Nollatutkimus Ei voida sanoa, että ryhmien välillä olisi eroa, koska sitä voi olla, jos voima olisi ollut riittävä	Onnenkantamoinen Voimaa sattui olemaan riittävästi ja ryhmien välillä on eroa.
Tehty (Ja ryhmäkoot täyttyvät)	Voimaa on riittävästi, mutta ryhmien välillä ei ole eroa	Voimaa on riittävästi ja ryhmien välillä on eroa

Tilastomenetelmät



Käyttötilanne	Vasteen tyyppi			
	Määrää mittaavaa (jatkuva muuttuja)			
	Symmetrinen jakauma	Hyvin vino jakauma	Luokkamuuttuja	Tapahtumaan kuluva aika
Jakauman kuvaaminen	Keskiarvo ja keskihajonta	Mediaani, kvartiiliväli, vaihteluväli	Suhteellinen osuus	Kaplan–Meier-menetelmä
Kahden riippumattoman ryhmän vertailu	Keskiarvojen vertailu ja riippumattomien otosten t-testi (<i>independent/unpaired t-test</i>)	Mediaanien vertailu ja Mann–Whitneyn U-testi	Ristiintaulukointi ja khi-neliötesti (tai Fisherin tarkka testi)	Kaplan–Meier-menetelmä ja logrank-testi
Usean riippumattoman ryhmän vertailu	Keskiarvojen vertailu ja varianssianalyysi (<i>analysis of variance, ANOVA</i>)	Mediaanien vertailu ja Kruskal–Wallis-testi	Ristiintaulukointi ja khi-neliötesti (tai Fisherin tarkka testi)	Kaplan–Meier-menetelmä ja logrank-testi
Kahden mittauskerran vertailu	Muutoksen keskiarvon vertailu ja toistettujen mittausten t-testi (<i>paired t-test</i>)	Muutoksen mediaanin tarkastelu ja Wilcoxonin testi	Suhteellisen osuuden muutoksen tarkastelu sekä McNemarin testi tai McNemar–Bowkerin testi	-
Useamman mittauskerran vertailu	Muutoksen tarkastelu ja toistettujen mittausten varianssianalyysi (<i>repeated measures ANOVA</i>)	Muutoksen tarkastelu ja Friedmanin testi (<i>Friedman test</i>)	Muutoksen tarkastelu ja Cochranin Q-testi	-
Muuttujien välisen riippuvuuden tutkiminen	Pearsonin korrelaatiokerroin (<i>Pearson’s correlation coefficient</i>)	Spearmanin korrelaatiokerroin	Ristiintaulukointi ja khi-neliötesti sekä suhteellinen riski (RR) tai vetosuhde (OR)	-
Vaihtelun selittäminen useilla selittävillä muuttujilla	Lineaarinen regressioanalyysi (<i>Linear regression analysis</i>)	Poissonin tai negatiivisen binomijakauman regressioanalyysi (<i>Poisson or negative binomial regression</i>)	Binäärinen tai multinomiaalinen logistinen regressioanalyysi (<i>binary or multinomial logistic regression</i>)	Coxin regressioanalyysi (<i>Cox regression</i>)

Julkaisuprosessin helpottamiseksi

- Tutkimussuunnitelma
 - Otokoko vaikuttaa resurssien määrään (esim. henkilöstö, talous, aika)
 - Mikä on kliinisesti merkitsevä ero?
- Eettiset luvat
 - Vaatii perustellun arvion otoskoosta
- Voimalaskelmat
 - Otokoon laskemiseksi ja voiman turvaamiseksi
- Protokolla-artikkeli
- Tutkimuksen rekisteröinti

Biostatistiikan neuvontapalvelut 1/2

- Neuvontapalvelut tarjoaa apua **tilastotieteellisissä kysymyksissä**:
 - tutkimuksen suunnittelussa, esim. otoskokolaskelmat, tutkimusasetelmat
 - tilastollisten analyysimenetelmien valinnassa ja analyysien toteutuksessa.
 - analyysien toteutuksessa
 - analyysien tulosten tulkinnassa
 - raportoinnissa, julkaisussa ja refereekommentteihin vastaamisessa
- Neuvontapalvelu on resursoitu aluksi somatiikan toimialueen tutkijoille
- Tutkimusten tulee tähdätä vertaisarvioituihin tutkimusartikkeleihin

Biostatistiikan neuvontapalvelut 2/2

- Opinnäytetöiden ohjaaminen on yliopiston vastuulla
- Kuvioiden tekeminen tutkijoiden vastuulla

Biostatistikot

- Pasi Ohtonen
 - Tytti Pokka
 - Marianne Haapea (50 %)
 - Marko Merikukka (20 %)
-
- Suurimman hyödyn neuvontapalvelusta saa, kun biostatistikko otetaan tutkimukseen mukaan jo tutkimuksen suunnitteluvaiheessa
 - Mikäli biostatistikon panos on ollut merkittävä julkaisun kannalta, hänet merkitään yhdeksi julkaisun tekijöistä

Yhteydenottolomake

- Työn kokonaismäärän hahmottamisen avuksi
- Töiden priorioimisen avuksi
- Dokumentointiin tilastollisen avun tarpeen määrästä Pohteella
- Lähetys osoitteeseen biostatistikot@ppshp.fi
- Verkkosivut: <https://oys.fi/tutkimus-ja-opetus/tutkijan-palvelut/tilastotieteen-palvelut/>

Yhteydenottolomake

1. Onko tutkimus?

- ☐ Osa lääketieteen syventäviä opintoja
- ☐ Osa väitöskirjatutkimusta
- ☐ Itsenäinen tutkimus

2. Tähtääkö tutkimus tieteelliseen vertaisarvioituun tutkimusartikkeliin?

- ☐ Kyllä
- ☐ Ei

3. Mikä on tutkimuksen tyyppi?

- ☐ Kliininen tutkimus
- ☐ Rekisteritutkimus
- ☐ Muu, mikä?
- ☐ Ei vielä tiedossa

4. Mikä on tutkimuksen vaihe?

- ☐ Tutkimuksen suunnittelu (tutkimusasetelma, otoskoko jne.)
- ☐ Aineistonkeruu (kyselylomakkeen suunnittelu, satunnaistaminen jne.)
- ☐ Aineiston tarkistaminen
- ☐ Analysointi
- ☐ Tulosten tulkinta ja raportointi
- ☐ Artikkelin referee-palautteeseen vastaaminen

5. Mikä on avun tarve?

- ☐ Otokoko ja voimalaskelmat
- ☐ Tilasto-ohjelman käytön opastus, minkä?
- ☐ Tilastomenetelmän valinta
- ☐ Tilastomenetelmän käyttö, minkä?
- ☐ Tulosten tulkinta
- ☐ Muu, mikä?

6. Missä yksikössä tutkimus suoritetaan?

Jos useita, niin kenellä on päävastuu?

7. Kuka on työnohjaaja?

Muuta hyödyllistä

Tilastotieteen palveluiden verkkosivuilla

- Mistä on hyvät taulukot tehty?
- Mistä on hyvät tilastograafit tehty?
- Verkkopohjainen kaaviotyökalu SmartDraw

Muulla verkossa

- SPSS-opas: <https://www.tutkijaportti.fi/syvarit/tilastolliset-menetelmat/>



Pohde

Pohjois-Pohjanmaan
hyvinvointialue